Finite Size Percolation in Regular Trees

Ery Arias-Castro

University of California, San Diego *

Abstract

In the context of percolation in a regular tree, we study the size of the largest cluster and the length of the longest run starting within the first d generations. As d tends to infinity, we prove almost sure and weak convergence results.

MSC 2010: Primary 60K35.
Keywords and Phrases: Percolation on trees; largest open cluster; longest success run; Galton-Watson processes; Chen-Stein method for Poisson approximation.

1 Introduction

Fix a positive integer r and let $\mathbb{T}$ be the infinite r-ary tree, rooted at ρ_0. We consider a Bernoulli percolation on $\mathbb{T}$. Formally, to each node $v \in \mathbb{T}$, we associate a random variable X_v, where the variables $\{X_v : v \in \mathbb{T}\}$ are i.i.d. Bernoulli with $\mathbb{P}(X_v = 1) = 1 - \mathbb{P}(X_v = 0) = p \in (0, 1)$. For a subset $A \subset \mathbb{T}$, let $X_A = \prod_{v \in A} X_v$. We say that A is *open* if $X_A = 1$.

1.1 The size of the largest cluster.

We use the term *cluster* to denote a connected component (i.e. subtree) of $\mathbb{T}$ when undirected. Let $\mathcal{K}$ denote the set of clusters in $\mathbb{T}$. For a node $v \in \mathbb{T}$, let $\mathrm{gen}(v)$ be its generation, i.e. the number of nodes in the shortest path from the root ρ_0 to v, not counting ρ_0. Note that $\mathrm{gen}(\rho_0) = 0$. Let $\mathbb{T}_d$ be the set of nodes with generation not exceeding d, namely $\mathbb{T}_d = \{v \in \mathbb{T} : \mathrm{gen}(v) \le d\}$. For a cluster $A \in \mathcal{K}$, we let $|A|$ denote its size (i.e. number of nodes) and $\rho(A)$ its root, namely $\rho(A) = \arg\min\{\mathrm{gen}(v) : v \in A\}$. For $d \in \mathbb{N}$, define K_d to be the size of the largest open cluster with root of generation not exceeding d:

$$K_d = \max\{|A| : A \in \mathcal{K},\, \rho(A) \in \mathbb{T}_d,\, X_A = 1\}.$$

In particular, K_0 is the size of the largest open cluster containing the root ρ_0.

In this paper we study the limit behavior of K_d, as $d \to \infty$. In the context of the one-dimensional lattice $\mathbb{Z}$, the corresponding results are often referred to as the Erdös-Rényi Law [8] and, in that context, our approach follows that of Arratia, Goldstein and Gordon [2]. In higher dimensions, the problem is much more intricate and many questions remain without answer. For a sample of sophisticated results, see e.g. [5, 16, 17]. The book by Grimmett [10] is a standard reference on percolation. For references more specific to trees, we refer the reader to a survey paper by Pemantle [15] and the book of Lyons and Peres [13]. Though the literature on percolation is vast, most of it focuses on the existence of an infinite cluster and its characteristics when it exists. On

*Department of Mathematics, University of California, San Diego {http://math.ucsd.edu/~eariasca}

the applications side, Patil and Taillie [14] identify regions of interest in a network by thresholding the response from each site in the network and computing connected components, which amounts to extracting the open clusters. One imagines that the largest cluster might receive the most attention. In particular, they mention monitoring water quality in a network of freshwater streams, where each stream may be modeled as a tree, though an irregular one.

It is well-known that, in the supercritical setting where $p > 1/r$, the cluster at the origin has positive probability of being infinite, and, in fact, $\mathbb{P}(K_d = \infty)$ tends to 1 as d increases. We restrict our attention to the subcritical and critical cases, i.e $p < 1/r$ and $p = 1/r$ respectively, where the cluster at the origin is finite with probability one. We start with the critical case, where we show that K_d behaves like the maximum of r^d independent random variables with distribution the total progeny of a Galton-Watson process with offspring distribution $\text{Bin}(r, 1/r)$. Let $\log_r$ denote the logarithm in base r.

Theorem 1. *Assume $p = 1/r$. Then with probability one,*

$$\frac{\log_r K_d}{d} \to 2, \quad d \to \infty.$$

Moreover,

$$\mathbb{P}(\log_r K_d \leq 2d + x) \to\to \exp(-C_1\, r^{-x/2}), \quad d \to \infty, \quad C_1 := \frac{2\, n^{-1/2}}{\sqrt{2\pi r(r-1)}}$$

For the subcritical case, we obtain similar results without transformation. Here, a Poisson approximation applies showing that K_d behaves like the maximum of $|\mathbb{T}_d| = r^{d+1}/(r-1)$ independent random variables with distribution the total progeny of a Galton-Watson process with offspring distribution $\text{Bin}(r, p)$. Define

$$\kappa = p(1-p)^{r-1}\ \frac{r^r}{(r-1)^{r-1}}. \tag{1}$$

Note that $\kappa < 1$ for all $p < 1/r$. Let $[x]$ denote the entire part of $x \in \mathbb{R}$.

Theorem 2. *Assume $p < 1/r$. Then with probability one,*

$$\frac{K_d}{d} \to \frac{1}{\log_r(1/\kappa)}, \quad d \to \infty.$$

Moreover, the sequence of random variables $(K_d - \mu_d : d \geq 0)$ is tight, where $\mu_d := \frac{d - \frac{3}{2}\log_r d}{\log_r(1/\kappa)}$. In addition, a subsequence $K_d - \mu_d$ converges weakly if, and only if, $a := \lim_{d\to\infty}(\mu_d - [\mu_d])$ exists, in which case the weak limit is $[Z + a] - a$, where

$$\mathbb{P}(Z \leq z) = \exp\left(-C_2\, \kappa^z\right),$$

for an explicit constant $C_2 > 0$ depending only on (p, r)

The behavior of the size of the largest open cluster in the subcritical regime is therefore similar in the context of the regular tree and in the context of the one-dimensional lattice, the latter corresponding to the length of the longest perfect head run in a sequence of coin tosses [2, Ex. 3].

1.2 The length of the longest run.

We use the term *run* for a path in $\mathbb{T}$ when directed away from the root ρ_0. Note that runs are special clusters. Let $\mathcal{R}$ denote the set of runs and define R_d to be the length of the longest open run with root of generation not exceeding d:

$$R_d = \max\{|A| : A \in \mathcal{R},\, \rho(A) \in \mathbb{T}_d,\, X_A = 1\}.$$

Of course, runs and clusters coincide in the one-dimensional lattice $\mathbb{Z}$. For a general reference on runs in dimension one, see [4]. Using the Chen-Stein method, Chen and Huo [6] proved results on the longest left-right run in a thin two-dimensional lattice of the form $([0,d] \times [0,a]) \cap \mathbb{Z}^2$, with the width a remaining constant. We also mention the work of Arias-Castro, Donoho and Huo [1] who used a statistic based on the longest run in a particular, non-planar graph to detect filaments in point-clouds.

The results we obtain for runs are parallel to those we obtain for clusters. In the critical case, we show that R_d behaves like the maximum of r^d independent random variables with distribution the height of a Galton-Watson process with offspring distribution $\mathrm{Bin}(r, 1/r)$.

Theorem 3. *Assume $p = 1/r$. Then with probability one,*

$$\frac{\log_r R_d}{d} \to 1, \quad d \to \infty.$$

Moreover, for any $x \in \mathbb{R}$,

$$\mathbb{P}\left(\log_r R_d \le d + x\right) \to \exp(-C_3\, r^{-x}), \quad d \to \infty, \quad C_3 := \frac{2rp}{r-1}.$$

In the subcritical case, we show that R_d behaves like the maximum of $|\mathbb{T}_d|$ independent random variables with distribution the height of a Galton-Watson process with offspring distribution $\mathrm{Bin}(r, p)$. Again, a Poisson approximation applies. The constant that appears in the exponent is only defined implicitly.

Theorem 4. *Assume $p < 1/r$. Then with probability one,*

$$\frac{R_d}{d} \to \frac{1}{\log_r(1/p) - 1}, \quad \text{as } d \to \infty.$$

Moreover, the sequence of random variables $(K_d - \nu_d : d \ge 0)$ is tight, where $\nu_d := \frac{d}{\log_r(1/p)-1}$. In addition, a subsequence $K_d - \nu_d$ converges weakly if, and only if, $a := \lim_{d\to\infty}(\nu_d - [\nu_d])$ exists, in which case the weak limit is $[Z + a] - a$, where

$$\mathbb{P}\left(Z \le z\right) = \exp\left(-C_4\, (rp)^z\right),$$

for an explicit constant $C_4 > 0$ depending only on (p, r)

1.3 Contents.

The rest of the paper is devoted to proving our results. In Section 2 we prove Theorem 1 and Theorem 2. In Section 3 we prove Theorem 3 and Theorem 4.

1.4 Additional Notation.

Let $\partial\mathbb{T}_d = \{v \in \mathbb{T} : \text{gen}(v) = d\}$. For a cluster A, let $\underline{A}$ denote the set of nodes not in A whose parents belong to A, and if $\rho(A) \neq \rho_0$, let $\mathring{A}$ denote the parent of $\rho(A)$. Also, define $(1-X)_A = \prod_{v\in A}(1-X_v)$. For two sequences of real numbers (a_n) and (b_n), we use the notation $a_n \sim b_n$ to indicate that $a_n/b_n \to 1$ and $a_n \asymp b_n$ to indicate that the ratio a_n/b_n is bounded away from zero and infinity, both understood as $n \to \infty$. Throughout the paper C denotes a finite, positive constant depending only on r and p, whose value may change with each appearance.

2 The size of the largest open cluster

In this section, we prove Theorem 1 and Theorem 2. We start with some notation. For a vertex $v \in \mathbb{T}$, let $K(v)$ be the size of the largest open cluster with root v,

$$K(v) = \max\{|A| : A \in \mathcal{K},\, \rho(A) = v,\, X_A = 1\}.$$

In particular,

$$K_d = \max\{K(v) : v \in \mathbb{T}_d\}.$$

The distribution of $K(v)$ does not depend on $v \in \mathbb{T}$, and, in fact, given $X_v = 1$, coincides with that of the total progeny of a Galton-Watson tree starting with one individual and with offspring distribution $\text{Bin}(r,p)$. Define

$$\psi_n = \mathbb{P}\left(K(v) = n\right), \quad \Psi_n = \mathbb{P}\left(K(v) > n\right).$$

Applying a well-known identity by Dwass [7] (called the Otter-Dwass formula in [13]), we get

$$\begin{aligned}
\psi_n &= \frac{p}{n}\mathbb{P}\left(\xi_1 + \cdots + \xi_n = n-1\right), \text{ where } \xi_1,\ldots,\xi_n \overset{\text{i.i.d.}}{\sim} \text{Bin}(r,p) \\
&= \frac{p}{n}\mathbb{P}\left(\text{Bin}(nr,p) = n-1\right) \\
&= \text{Cat}_n\, p^n (1-p)^{n(r-1)+1},
\end{aligned}$$

where

$$\text{Cat}_n := \frac{1}{n}\binom{nr}{n-1} = \frac{1}{(r-1)n+1}\binom{rn}{n}$$

is the *nth generalized Catalan number* [11], which among other interpretations, is the number of subtrees of $\mathbb{T}$ of size n rooted at the origin, i.e.

$$\text{Cat}_n = |\{A \in \mathcal{K} : \rho(A) = \rho_0,\, |A| = n\}|.$$

We could have obtained the expression for ψ_n using this definition of Cat_n. Indeed, for $n > 0$, $K(v) = n$ if, and only if, there is a (unique) subtree A with $|A| = n$, $\rho(A) = v$ and $X_A(1-X)_{\underline{A}} = 1$, so that A cannot be extended and still be an open cluster. We then use the fact that a subtree of size n has exactly $(r-1)n+1$ children. With the use of Stirling's formula, we arrive at the following conclusions; see also [3, 12].

Lemma 1. *In the critical case $p = 1/r$,*

$$\Psi_n \sim \frac{C_1}{\sqrt{n}}.$$

In the subcritical case $p < 1/r$,

$$\Psi_n \sim C_5\,\frac{\kappa^{n+1}}{n^{3/2}}, \quad C_5 := \frac{1}{\sqrt{2\pi}(1-\kappa)}\,\frac{(1-p)r^{1/2}}{(r-1)^{3/2}}.$$

2.1 Proof of Theorem 1

Define

$$K_d^\partial := \max\{K(v) : v \in \partial\mathbb{T}_d\}.$$

We first prove that the conclusions of Theorem 1 hold for K_d^∂. For $x \in \mathbb{R}$, let $n_d(x) = [r^{2d+x}]$. As K_d^∂ only involves independent random variables, we have

$$\mathbb{P}\left(\log_r K_d^\partial \le 2d + x\right) = \mathbb{P}\left(K_d^\partial \le n_d(x)\right) = (1 - \Psi_{n_d(x)})^{r^d} = \exp(-C_1\, r^{-x/2} + O(r^{-d-x})).$$

Letting $d \to \infty$, we obtain the weak convergence, and by choosing $x = \varepsilon d$, with $\varepsilon > -2$ fixed, and applying the Borel-Cantelli Lemma, we obtain the almost sure convergence.

It therefore suffices to show that $K_d = (1 + o_P(1))K_d^\partial$. Clearly, $K_d \ge K_d^\partial$, so we focus on the upper bound. Define

$$B_d = \{v \in \partial\mathbb{T}_d : K(v) > r^d/d\}, \quad B_d^2 = \{v \in \partial\mathbb{T}_d : K(v) > r^{2d}/d\}.$$

For any open cluster A with $\rho(A) \in \mathbb{T}_d$, we have

$$|A| = |A \cap \mathbb{T}_{d-1}| + \sum_{v \in A \cap \partial\mathbb{T}_d} K(v) \le r^d + r^{2d}/d + \sum_{v \in A \cap B_d} K(v).$$

We turn to bounding the sum. We first show that, with probability tending to one, there is no open cluster A containing three or more nodes in B_d. Indeed, take $v_1, v_2, v_3 \in \partial\mathbb{T}_d$ distinct. Let w denote their most recent common ancestor and let $k = d - \text{gen}(w)$. Either the paths $v_j \to \rho_0$ meet at w for the first time or two of the paths meet at a node u with $\text{gen}(u) > \text{gen}(w)$, in which case we let $\ell = d - \text{gen}(u)$. Now, the nodes v_1, v_2, v_3 belong to the same open cluster if, and only if, the smallest subtree containing w and v_1, v_2, v_3 is open, and this subtree is of size $\ell + 2k + 1$, and therefore, the probability that they belong to the same open cluster is $p^{\ell+2k+1}$. In addition, the number of such triplets is bounded by

$$\left(r^{d-k}\binom{r^k}{3}\right) \cdot \left(r^k r^{k-\ell}\binom{r^\ell}{2}\right)\binom{r^k}{3}^{-1} \asymp r^{d+k+\ell}.$$

The first factor comes from the fact that the three nodes are leaves of a subtree with root at generation $d - k$. Given that, the second factor comes from the fact that two of them belong to a subtree of that subtree with root at (relative) generation $k - \ell$. Hence, remembering that $p = 1/r$ and using Lemma 1, we have

$$\begin{aligned}
\mathbb{P}\left(\exists A \in \mathcal{K} : X_A = 1,\, |A \cap B_d| \ge 3\right) &\le C\,\mathbb{P}\left(K(v) > r^d/d\right)^3 \cdot \sum_{k=0}^{d}\sum_{\ell=0}^{k} r^{d+k+\ell} p^{\ell+2k+1} \\
&\le C\,(r^d/d)^{-3/2} r^d \asymp d^{3/2} r^{-d/2}.
\end{aligned}$$

By the same token, with probability tending to one (in fact of order at most d/r^d), there is no open cluster A containing two or more nodes in B_d^2. Now, when $|A \cap B_d| \le 2$ and $|A \cap B_d^2| \le 1$, we have

$$\sum_{v \in A \cap B_d} K(v) \le \max_{v \in A \cap B_d} K(v) + r^{2d}/d \le K_d^\partial + r^{2d}/d.$$

In the end, with probability tending to one,

$$|A| \le r^d + 2\, r^{2d}/d + K_d^\partial,$$

for any open cluster A with $\rho(A) \in \mathbb{T}_d$. Hence,

$$K_d \le K_d^\partial + O_P(r^{2d}/d),$$

and we conclude by the fact that K_d^∂ is of order exceeding r^{2d}/d with probability tending to one.

2.2 Proof of Theorem 2

The proof of the almost sure convergence may be obtained following the arguments provided in Section 2.1 or using the bounds we are about to prove below. We omit details.

The proof of the weak convergence is based on the Chen-Stein method for Poisson approximation as formulated by Arratia, Goldstein and Gordon [2]. Define

$$Y_A = \begin{cases} X_A(1-X)_{\underline{A}}, & \rho(A) = \rho_0, \\ X_A(1-X)_{\mathring{A}}(1-X)_{\underline{A}}, & \rho(A) \neq \rho_0; \end{cases}$$

Also, let $\mathcal{K}_{d,n}$ be the set of clusters of size exceeding n with root in $\mathbb{T}_d$, and define

$$W_{d,n} = \sum_{A \in \mathcal{K}_{d,n}} Y_A.$$

By definition,

$$\{K_d \leq n\} = \{Y_A = 0, \forall A \in \mathcal{K}_{d,n}\} = \{W_{d,n} = 0\}.$$

We approximate the law of $W_{d,n}$ by the Poisson distribution with same mean $\lambda_{d,n} = \mathbb{E}(W_{d,n})$. We start by estimating $\lambda_{d,n}$ using Lemma 1, obtaining

$$\begin{aligned} \lambda_{d,n} &= \sum_{A \in \mathcal{K}_{d,n}} \mathbb{P}(Y_A = 1) \\ &= \mathbb{P}(K(\rho_0) > n) + (1-p) \sum_{v \in \mathbb{T}_d, v \neq \rho_0} \mathbb{P}(K(v) > n) \\ &= \Psi_n + (1-p)(|\mathbb{T}_d| - 1)\Psi_n. \end{aligned}$$

In particular, as $n, d \to \infty$,

$$\lambda_{d,n} \sim C_2\, r^d n^{-3/2} \kappa^{n+1}, \quad C_2 := \frac{C_5(1-p)r}{r-1}.$$

For a cluster $A \in \mathcal{K}_{d,n}$, define its neighborhood $\mathcal{B}(A)$ as the set of clusters $B \in \mathcal{K}_{d,n}$ such that

$$(\mathring{B} \cup B \cup \underline{B}) \ \cap \ (\mathring{A} \cup A \cup \underline{A}) \neq \emptyset.$$

Define the following sums

$$\begin{aligned} F_{d,n} &= \sum_{A \in \mathcal{K}_{d,n}} \sum_{B \in \mathcal{B}(A)} \mathbb{P}(Y_A = 1)\, \mathbb{P}(Y_B = 1), \\ G_{d,n} &= \sum_{A \in \mathcal{K}_{d,n}} \sum_{B \in \mathcal{B}(A), B \neq A} \mathbb{P}(Y_A = Y_B = 1), \\ H_{d,n} &= \sum_{A \in \mathcal{K}_{d,n}} \mathbb{E}\left(\left|\mathbb{E}\left(Y_A - \mathbb{E}(Y_A)\,|Y_B, B \notin \mathcal{B}(A)\right)\right|\right). \end{aligned}$$

Then by the second part of [2, Th. 1],

$$|\mathbb{P}(W_{d,n} = 0) - \exp(-\lambda_{d,n})| \leq F_{d,n} + G_{d,n} + H_{d,n}.$$

For $x \in \mathbb{R}$, define $n_d(x) = [\mu_d + x]$. When x is fixed and $d \to \infty$, $\lambda_{d,n_d(x)} \asymp 1$, with

$$\lambda_{d,n_d(x)} \to C_2\, \kappa^{[a+x]-a+1}, \ \ \text{when } \mu_d - [\mu_d] \to a, \ x - [x] \neq 1 - a,$$

with

$$\mathbb{P}\left([Z+a]-a\le x\right)=\exp(-C_2\,\kappa^{[a+x]-a+1}).$$

Therefore, to conclude it suffices to prove that $F_{d,n}, G_{d,n}, H_{d,n}\to 0$ when $d,n\to\infty$ in such a way that $\lambda_{d,n}\asymp|\mathbb{T}_d|\Psi_n\asymp 1$. First, $H_{d,n}=0$ by independence of Y_A and $Y_B, B\notin\mathcal{B}(A)$. For $G_{d,n}$, the only pairs $A,B\in\mathcal{K}_{d,n}$ that contribute to the sum satisfy either $\mathring{B}\in\underline{A}$ or $\mathring{A}\in\underline{B}$, and in both cases

$$\mathbb{P}\left(Y_A=Y_B=1\right)=(1-p)^{-1}\mathbb{P}\left(Y_A=1\right)\mathbb{P}\left(Y_B=1\right).$$

Hence, using the fact that there are Cat_m subtrees of size m with a given root, each with $(r-1)m+1$ children, and then Lemma 1, we have

$$\begin{aligned}G_{d,n}&\le 2(1-p)^{-1}|\mathbb{T}_d|\sum_{m>n}\mathrm{Cat}_m\,p^m(1-p)^{(r-1)m+1}((r-1)m+1)\cdot\Psi_n\\&\le C\,\lambda\sum_{m>n}m\psi_m=C\,\lambda\left((n+1)\Psi_n+\sum_{m>n}\Psi_m\right)\asymp n^{-1/2}\kappa^n\to 0,\quad n\to\infty.\end{aligned}$$

For $F_{d,n}$, the only pairs $A,B\in\mathcal{K}_{d,n}$ that contribute to the sum satisfy either $\mathring{B}\in\mathring{A}\cup A\cup\underline{A}$ or $\mathring{A}\in\mathring{B}\cup B\cup\underline{B}$. The computations are then similar.

3 The length of the longest open run

The arguments are parallel to those provided in Section 2. For $A\subset\mathbb{T}$, define its height as $\tau(A)=\sup\{\mathrm{gen}(v):v\in A\}-\mathrm{gen}(\rho(A))$. For a vertex $v\in\mathbb{T}$, let $R(v)$ be the length of the longest run with root v,

$$R(v)=1+\max\{\tau(A):A\in\mathcal{K},\rho(A)=v\}.$$

In particular,

$$R_d=\max\{R(v):v\in\mathbb{T}_d\}.$$

The distribution of $R(v)$ does not depend on $v\in\mathbb{T}$, and, in fact, given $X_v=1$, coincides with that of the height (plus one), i.e. extinction time, of a Galton-Watson tree with offspring distribution $\mathrm{Bin}(r,p)$. Define

$$\phi_h=\mathbb{P}\left(R(v)=h\right),\quad \Phi_h=\mathbb{P}\left(R(v)>h\right).$$

We have the following results on the asymptotic behavior of Φ_h [3].

Lemma 2. *In the critical case $p=1/r$,*

$$\Phi_h\sim\frac{C_3}{h}.$$

In the subcritical case $p<1/r$, there is an implicit constant $C_6>0$ such that

$$\Phi_h\sim C_6\,(rp)^h.$$

Let $\mathrm{Cat}_{n,h}$ denote the number of subtrees rooted at the origin, of size n and height h. See [9] for some results on $\mathrm{Cat}_{n,h}$. As in Section 2, we can argue that

$$\phi_h=\sum_{n>h}\mathrm{Cat}_{n,h}\,p^n(1-p)^{(r-1)n+1},\ \text{ implying }\ \Phi_h=\sum_{\ell>h}\sum_{n>\ell}\mathrm{Cat}_{n,\ell}\,p^n(1-p)^{(r-1)n+1}.$$

3.1 Proof of Theorem 3

The proof is based on the following observation

$$R_d^\partial \le R_d \le R_d^\partial + d, \quad R_d^\partial := \max\{R(v) : v \in \partial\mathbb{T}_d\},$$

where the d term bounds the length of any run in $\mathbb{T}_{d-1}$. As R_d^∂ only involves independent random variables,

$$\mathbb{P}\left(R_d^\partial \le h\right) = (1-\Phi_h)^{r^d}. \tag{2}$$

Choosing $h = r^{[d/2]}$ and using Lemma 2, we obtain

$$\mathbb{P}\left(R_d^\partial \le r^{[d/2]}\right) \le \exp(-C\, r^{-d/2}), \quad \text{for some } C > 0,$$

so that, applying the Borel-Cantelli Lemma, $R_d^\partial \ge r^{d/2}$ eventually, with probability one. Hence, $\log_r R_d = (1 + o_P(1)) \log_r R_d^\partial$, and it is therefore enough to prove the results for R_d^∂ in place of R_d. The almost sure convergence is obtained in a similar way by choosing $h = r^{(1+\varepsilon)d}$ with ε fixed, either positive or negative. For the weak convergence, fix x and let $h_d(x) = [r^{d+x}]$. By Lemma 2 and (2), we have

$$\mathbb{P}\left(\log_r R_d^\partial \le d + x\right) \to \exp(-C_3\, r^{-x}), \quad d \to \infty.$$

3.2 Proof of Theorem 4

We again omit the details of the proof of the almost sure convergence and focus on proving the weak convergence. Let $\mathcal{K}_{d,h}$ denote the set of clusters with root in $\mathbb{T}_d$ and height exceeding h. We use the notation introduced in Section 2.2, with $\mathcal{K}_{d,h}$ in place of $\mathcal{K}_{d,n}$. By definition,

$$\{R_d \le h\} = \{W_{d,h} = 0\}.$$

Using Lemma 2, we obtain

$$\begin{aligned}
\lambda_{d,h} &= \sum_{A \in \mathcal{K}_{d,h}} \mathbb{P}\left(Y_A = 1\right) \\
&= \mathbb{P}\left(R(\rho_0) > h\right) + (1-p) \sum_{v \in \mathbb{T}_d} \mathbb{P}\left(R(v) > h\right) \\
&= \Phi_h + (1-p)(|\mathbb{T}_d| - 1)\Phi_h.
\end{aligned}$$

In particular, as $h, d \to \infty$,

$$\lambda_{d,h} \sim C_4\, r^d (rp)^{h+1}, \quad C_4 := \frac{C_6\,(1-p)}{p(r-1)}.$$

For $x \in \mathbb{R}$, define $h_d(x) = [\nu_d + x]$. When x is fixed and $d \to \infty$, we have $\lambda_{d,h_d(x)} \asymp 1$, with

$$\lambda_{d,h_d(x)} \to C_4 (rp)^{[a+x]-a+1}, \quad \text{when } \nu_d - [\nu_d] \to a,\ x - [x] \ne 1 - a.$$

It then suffices to show that $F_{d,h}, G_{d,h}, H_{d,h} \to 0$ when $d, h \to \infty$ in such a way that $\lambda_{d,h} \asymp |\mathbb{T}_d|\Phi_h \asymp 1$, and the computations are parallel to those in Section 2.2. We focus on $G_{d,h}$. Fix $\tilde{p} \in (p, 1/r)$

and let $\tilde{\Phi}_h$ be defined as Φ_h, with $\tilde{p}$ in place of p. For h large enough, we then have

$$\begin{aligned}
G_{d,h} &\leq 2 \sum_{A \in \mathcal{K}_{d,h}} \sum_{\mathcal{B}(A), \mathring{B} \in \underline{A}} (1-p)^{-1} \mathbb{P}(Y_A = 1) \, \mathbb{P}(Y_B = 1) \\
&\leq C \, |\mathbb{T}_d| \sum_{\ell > h} \sum_{n > \ell} \mathrm{Cat}_{n,\ell} \, p^n (1-p)^{(r-1)n+1} ((r-1)n+1) \cdot \Phi_h \\
&\leq C \, \lambda \sum_{\ell > h} \sum_{n > \ell} \mathrm{Cat}_{n,\ell} \, \tilde{p}^n (1-\tilde{p})^{(r-1)n+1} \\
&= C \, \lambda \, \tilde{\Phi}_h \asymp (r\tilde{p})^h \to 0, \quad d \to \infty.
\end{aligned}$$

Acknowledgements

The author would like to thank Philippe Flajolet for fruitful conversations and Jason Schweinsberg for reading an early version of the manuscript, pointing out some errors and helping with the proof of Theorem 1. This work was partially supported by a grant from the National Science Foundation (DMS-0603890) and a grant from the Office of Naval Research (N00014-09-1-0258).

References

[1] E. Arias-Castro, D. L. Donoho, and X. Huo. Adaptive multiscale detection of filamentary structures in a background of uniform random points. *Ann. Statist.*, 34(1):326–349, 2006.

[2] R. Arratia, L. Goldstein, and L. Gordon. Two moments suffice for Poisson approximations: the Chen-Stein method. *Ann. Probab.*, 17(1):9–25, 1989.

[3] K. B. Athreya and P. E. Ney. *Branching processes*. Springer-Verlag, New York, 1972. Die Grundlehren der mathematischen Wissenschaften, Band 196.

[4] N. Balakrishnan and M. V. Koutras. *Runs and scans with applications*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], New York, 2002.

[5] C. Borgs, J. T. Chayes, H. Kesten, and J. Spencer. The birth of the infinite cluster: finite-size scaling in percolation. *Comm. Math. Phys.*, 224(1):153–204, 2001. Dedicated to Joel L. Lebowitz.

[6] J. Chen and X. Huo. Distribution of the length of the longest significance run on a Bernoulli net and its applications. *J. Amer. Statist. Assoc.*, 101(473):321–331, 2006.

[7] M. Dwass. The total progeny in a branching process and a related random walk. *J. Appl. Probability*, 6:682–686, 1969.

[8] P. Erdős and A. Rényi. On a new law of large numbers. *J. Analyse Math.*, 23:103–111, 1970.

[9] P. Flajolet, Z. Gao, A. Odlyzko, and B. Richmond. The distribution of heights of binary trees and other simple trees. *Combin. Probab. Comput.*, 2(2):145–156, 1993.

[10] G. Grimmett. *Percolation*, volume 321 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, second edition, 1999.

[11] P. Hilton and J. Pedersen. Catalan numbers, their generalization, and their uses. *Math. Intelligencer*, 13(2):64–75, 1991.

[12] D. P. Kennedy. The Galton-Watson process conditioned on the total progeny. *J. Appl. Probability*, 12(4):800–806, 1975.

[13] R. Lyons and Y. Peres. *Probability on trees and networks*. Cambridge University Press, 2010. In preparation. Current version available at http://mypage.iu.edu/~rdlyons/.

[14] G. P. Patil and C. Taillie. Geographic and network surveillance via scan statistics for critical area detection. *Statist. Sci.*, 18(4):457–465, 2003.

[15] R. Pemantle. Tree-indexed processes. *Statist. Sci.*, 10(2):200–213, 1995.

[16] M. D. Penrose. A central limit theorem with applications to percolation, epidemics and Boolean models. *Ann. Probab.*, 29(4):1515–1546, 2001.

[17] M. D. Penrose and A. Pisztora. Large deviations for discrete and continuous percolation. *Adv. in Appl. Probab.*, 28(1):29–52, 1996.